Bożena Czernecka-Rej

"Nieobliczalna obliczalność", Maria Piesko, Kraków 2011 : [recenzja]

Studia Philosophiae Christianae 48/1, 214-222

2012

35. Nonsens: „Pole relacji oznaczamy symbolem P(R): P(R) = D(R) + D'(R)" (s. 147). Znak: + (nigdzie wcześniej nie zdefiniowany) jest użyty nonsensownie. Powinien być użyty znak sumy zbiorów: ∨ zdefiniowany na s. 143.
35. Nonsens: „Relacje zachodzące między elementami określonego zbioru można ocenić pod pewnymi szczególnych własności, jakie one wykazują. W ten sposób wyróżniono następujące relacje:
    A. Relacja zwrotna (refleksyjna)
    ∧ x xRx" (s. 148).
    Pomijamy błąd gramatyczny. Termin „refleksyjna" nie jest używany (raczej „refleksywna"). Definicja jednej z formalnych własności relacji dwuargumentowych – zwrotności jest nieskładna: brakuje definiendum, funktora definicyjnego oraz relatywizacji do danego zbioru. W taki sam sposób są nonsensownie sformułowane definicje wszystkich pozostałych własności – s. 148-149.

Sporządzona lista uwag nie jest kompletna ale, jak sądzę, wystarczająca, by uzasadnić moją negatywną opinię o recenzowanej pozycji. W swoim obecnym kształcie książka prof. K. Pawłowskiego nie powinna być w ogóle opublikowana. Pozostaje mi tylko żywić nadzieję, że w przyszłości nie powtórzą się tego rodzaju niefortunne publikacje firmowane przez naszą Uczelnię. Fakt ukazania się tej książki rzuca cień na prestiż całego naszego Uniwersytetu.

Kordula Świętorzecka
Instytut Filozofii UKSW

Maria Piesko, *Nieobliczalna obliczalność*, Konsorcjum Akademickie, Kraków 2011, ss. 216.

Na progu drugiej dekady XXI w. nikogo nie trzeba przekonywać, że żyjemy w epoce komputerów. Powierzamy komputerom coraz więcej

zadań, ich rola rośnie w wielu dziedzinach naszego życia. Fakt ten skłania do stawiania pytań o to, czy komputer to rzeczywiście wielce użyteczna maszyna działająca według określonego algorytmu, która może adekwatnie symulować ludzkie myślenie. Innymi słowy, czy możliwa jest tzw. sztuczna inteligencja (od angielskiej nazwy *Artificial Intelligence*, wprowadzonej w 1956 r. przez Johna McCarthy'ego, powszechnie używa się skrótu AI), czy dziś jest ona czymś więcej niż tylko tworem *science-fiction*?

W badaniach nad AI fundamentalne znaczenie posiada refleksja nad naturą obliczalności. Zainteresowanie uczonych pojęciem obliczalności datuje się od wystąpienia Davida Hilberta na Międzynarodowym Kongresie Matematycznym (w 1900 r.), na którym przedstawił on – jako wyzwanie dla matematyków zaczynającego się stulecia – listę 23 najważniejszych nierozstrzygniętych problemów matematycznych. Dziesiąty problem Hilberta, tzw. Entscheidungsproblem, sprowadza się do pytania: czy istnieje ogólna, algorytmiczna procedura rozstrzygania zagadnień matematycznych należących do pewnej dobrze zdefiniowanej klasy? Licząc na pozytywną odpowiedź, Hilbert opracował program zbudowania matematyki na niepodważalnych podstawach. Program ten, zwany matematycznym formalizmem, obaliły wyniki Kurta Gödla z 1931 r., dzięki którym wiemy, że w matematyce zawsze będzie pewne *ignorabimus*.

Zagadnienia związane z obliczalnością *explicite* odwołują się do rozstrzygnięć brytyjskiego matematyka, Alana Turinga, który chciał rozwiązać wspomniany problem Hilberta. W artykule z 1936 r. *On Computable Numbers with an Application to the Entscheidungsproblem* (*O liczbach obliczalnych z ich zastosowaniem dla problemu rozstrzygalności*) wykorzystał on pojęcie abstrakcyjnej maszyny obliczeniowej, która wykonuje proste operacje na napisach według ściśle ustalonych reguł. Maszyna ta, nazwana przez Alonzo Churcha *maszyną Turinga* (w skrócie: TM), dostarcza funkcjonalnej (operacyjnej) definicji obliczeń, algorytmu, procedury mechanicznej: algorytm jest tym, co może być wykonywane przez TM. Problem Hilberta wyrażony w języku Turinga – tzw. „problem stopu" – wyraża się w pytaniu: czy istnieje uni-

wersalny algorytm, który rozstrzygałby o zakończeniu pracy dowolnej TM? Turing wykazał, że jest on niemożliwy do rozstrzygnięcia z powodu istnienia nieobliczalnej liczby rzeczywistej, czyli że nie istnieje żadna taka procedura (żaden algorytm, żaden program komputerowy), która umożliwiałaby realizację oczekiwań Hilberta. Innymi słowy, nie ma algorytmu, który umożliwiałby ustalenie, czy *n*-ty program komputerowy kiedykolwiek wygeneruje *n*-tą cyfrę (i zakończy pracę), czy też nigdy jej nie wygeneruje i wobec tego nigdy się nie zatrzyma (nie zakończy pracy).

Od czasu konstrukcji pierwszych komputerów i pierwszych rozważań o sztucznej inteligencji TM stanowi punkt odniesienia większości prac na temat „intelektualnych" możliwości komputerów. Współcześni uczeni poszukują odpowiedzi na następujące pytania: czy TM może być adekwatnym modelem umysłu i działania ludzkiego?; czy może ona modelować całość umysłu, czy tylko pewne jego aspekty, określone operacje umysłowe?; jakie inne jeszcze maszynowe modele (teoretyczne lub praktyczne) mogą być pomocne w opisie poszczególnych rodzajów myślenia i działania człowieka? Podkreśla się w literaturze, że pojęcie obliczalności stanowi swoistego rodzaju „broń obosieczną". Z jednej strony jest ono wykorzystywane jako podstawa postulowania analogii pomiędzy sztucznymi a naturalnymi podmiotami poznawczymi, z drugiej zaś w oparciu o to pojęcie próbuje się wykazać nieprawomocność komputacyjnych podejść do badania umysłu.

O powyższych zagadnieniach traktuje książka Marii Piesko *Nieobliczalna obliczalność*, wydana w tym roku przez Konsorcjum Akademickie (Wydawnictwo WSE w Krakowie, WSIiZ w Rzeszowie i WSZiA w Zamościu). Autorka jest doktorem filozofii, współpracuje z Ośrodkiem Badań Interdyscyplinarnych w Krakowie.

Praca składa się z pięciu rozdziałów, wstępu, zakończenia i spisu wykorzystanej literatury. Szkoda, że Autorka nie sporządziła indeksu (osobowego i rzeczowego), który byłby wielce użyteczny, zwłaszcza przy wybiórczej lekturze książki. W dwóch pierwszych rozdziałach Piesko podejmuje analizę zagadnienia obliczalności z perspektywy historycznej, przedstawia szeroki kontekst powstania teorii obliczalności i jej

pierwotne zastosowania. W trzech pozostałych uzasadnia tezę o zachodzeniu procesów (teoretycznie i praktycznie) nieobliczalnych, w związku z którą rozważa w sposób systematyczny możliwość realizacji programu AI. Autorka – za Turingiem – nie aprobuje postawy „chowania głowy w piasek", która opiera się na przekonaniu, że „byłoby to zbyt okropne, gdyby jakakolwiek maszyna miała posiąść zdolność, którą *homo sapiens* chlubi się najbardziej" (s. 9). Nie przesądza wprawdzie, że realizacja programu AI jest możliwa, ale też nie uważa, że spekulacje na ten temat należą wyłącznie do zakresu fantastyki naukowej. Podejmuje analizę możliwości, a zwłaszcza ograniczeń maszyn cyfrowych, stających na drodze do zaprojektowania „sztucznego mózgu".

W rozdziale pierwszym, najobszerniejszym w całej książce, zgodnie z tytułem Autorka wprowadza czytelnika *W krainę abstrakcji*. Opisuje główne nurty refleksji metamatematycznej oraz ważne fakty w rozwoju samej matematyki XIX/XX w., takie jak: odkrycie geometrii nieeuklidesowych, unifikację (arytmetyzację) matematyki i aksjomatyzację arytmetyki liczb naturalnych, odkrycie fundamentów matematyki w postaci teorii mnogości, a następnie antynomii na jej gruncie, powstanie trzech szkół w zakresie podstaw matematyki (logicyzmu, intuicjonizmu i formalizmu), poszukujących rozwiązania problemu antynomii, powstanie logiki współczesnej, udowodnienie twierdzeń limitacyjnych przez Gödla. Przy każdej okazji stara się znaleźć związki owych faktów z ideami zawartymi w artykule Turinga (z 1936 r.), który odegrał istotną rolę w kształtowaniu się koncepcji komuptacjonistycznych.

Pracy Turinga poświęcony jest drugi rozdział zatytułowany *On Computable Numbers*. Piesko prezentuje w nim jeden z formalizmów ujmujących w ścisłą postać intuicyjne rozumienie obliczalności – teorię TM. Podaje dowód nierozstrzygalności problemu stopu oraz oparty na nim dowód nierozstrzygalności arytmetyki. Przytacza argumentację na rzecz ważnej z punktu widzenia problematyki podstaw matematyki oraz AI tezy Turinga, stwierdzającej adekwatność zaproponowanej przez Turinga definicji obliczalności. Chociaż problem AI nie występuje wprost w artykule z 1936 r., niemniej jego autor określa dokładnie, co może, a czego nie może dokonać maszyna licząca, którą najczęściej

bierze się pod uwagę, szukając sposobów imitacji ludzkiej inteligencji. Zaproponowany przez Turinga formalizm pozwala na wyznaczenie zarówno granic możliwości maszyn (problem stopu), jak i ich „siły obliczeniowej". Piesko przestrzega jednak przed wykorzystywaniem twierdzenia Turinga o nierozstrzygalności w argumentacji przeciwko AI, uważając, że twierdzenie to wyznacza przede wszystkim granice ludzkiej metodycznej działalności. Z drugiej strony, z tego samego powodu, bezpodstawne jest powoływanie się na nie w argumentacji na rzecz AI, tak jak problematyczne jest utożsamianie inteligencji z efektywnym obliczaniem.

Rozdział trzeci stanowi wprowadzenie do drugiej części pracy, będącej próbą zrozumienia, na czym polega nieobliczalność. Autorka omawia tu różne wersje tezy Turinga, częściej nazywanej tezą Churcha lub Churcha-Turinga. Teza ta (w oryginalnym sformułowaniu) głosi, że każda funkcja, która może zostać obliczona przez człowieka wyposażonego wyłącznie w papier i ołówek oraz postępującego w mechaniczny sposób, jest obliczalna za pomocą TM. Inaczej: każda funkcja efektywnie obliczalna jest rekurencyjna (obliczalna za pomocą TM). Autorka analizuje jej tzw. „tezy potomne": tezę M – „cokolwiek może być obliczone przez maszynę (...), jest obliczalne przez TM" (s. 119), tezę S – „każdy proces, dla którego można podać opis matematyczny (...), może być symulowany przez TM" (s. 120), tezę B – „zachowanie żywych organizmów może być symulowane za pomocą TM" (s. 121) oraz tezę o AI – „procesy mentalne mogą być symulowane za pomocą TM" (s. 121). Tezy te są uporządkowane w ten sposób, że każda następna jest silniejsza – od jej oryginalnego sformułowania w pismach Turinga i Churcha: od tezy M dotyczącej maszyn, przez tezę S odnoszącą się do procesów fizycznych, tezę B uwzględniającą specyfikę żywych organizmów, aż do tezy o AI. Według Piesko, jedynie pierwsza z tych tez została dobrze uzasadniona przez Turinga, tezy potomne zaś wymagają dodatkowych argumentów.

O tym, co nieobliczalne (teoretycznie) traktuje rozdział czwarty. Funkcji nieobliczalnych jest nieskończenie wiele więcej niż obliczalnych. Maszyn TM jest jedynie przeliczalnie nieskończenie wiele, a nieskończonych ciągów liczb (nawet nieskończonych ciągów 0-1-owych) jest

nieprzeliczalnie wiele. Piesko wskazuje także „miejsca dla nieobliczalności" we współczesnych teoriach fizycznych (np. procesy analogowe i związane z nimi operowanie liczbami rzeczywistymi niewymiernymi, teoria kwantów, ogólna teoria względności), zastrzegając, iż przy przejściu na teren fizyki zmienia się nie tylko perspektywa badań, ale też samo pojęcie (nie)obliczalności.

W ostatnim rozdziale Autorka koncentruje się na nieobliczalności praktycznej, zmieniając po raz kolejny perspektywę rozważań. Jej zdaniem, teoria obliczalności, wykorzystująca formalizm zwykłych TM, posiada ograniczone zastosowanie praktyczne. Model TM jest istotny, aczkolwiek niewystarczający do opisu działania współczesnych systemów komputerowych, w których coraz większą rolę odgrywa interakcja urządzeń ze środowiskiem i urządzeń pomiędzy sobą. Piesko proponuje zastąpić ten model teorią maszyn interaktywnych. Skoro bowiem człowiek działa w świecie w bardzo wielu relacjach, uczy się wymieniając informacje z otoczeniem, nie inaczej powinno być z maszyną, aspirującą do miana inteligentnej. W „nieobliczalnym środowisku" – a taki jest, jak stara się wykazać Autorka, fizyczny świat – maszyna interakcyjna jest w istotny sposób „silniejsza" od zwykłej TM. Jak dotąd, nie został jednak opracowany model matematyczny maszyn interaktywnych.

W końcowej partii recenzowanej książki Piesko pisze, iż nie chce być postrzegana jako zbytnia optymistka w kwestii AI, traktująca interaktywne TM jako cudowny środek do zbudowania sztucznego mózgu. Nie twierdzi ona, że można zrekonstruować ludzkie myślenie za pomocą komputerów czy jakichkolwiek innych sztucznych środków. Uważa jednak, że teza o zasadniczej wykonalności tego zadania winna należeć do metodologicznych założeń wszelkich badań nad AI.

Intrygujący jest sam tytuł książki. Obydwa tytułowe słowa zostały użyte w różnych znaczeniach: przymiotnik „nieobliczalna" w sensie potocznym, natomiast rzeczownik „obliczalność" w znaczeniu technicznym. Według takiej interpretacji, tytuł recenzowanej pozycji ma znaczyć tyle, że matematyczna obliczalność jest nieprzewidywalna, wymykająca się spod kontroli, a w szczególności to, że znajduje się ona poza zasięgiem dostępnych formalizmów, w tym poza zasięgiem TM.

Warto zauważyć, że na temat natury ludzkiego umysłu i jego matematycznego modelu w kontekście pytania o możliwość stworzenia sztucznej inteligencji, wypowiedział się również sam Turing. W artykule z 1950 r. *Maszyny liczące a inteligencja*, stanowiącym deklarację programową AI, przyznał, że pytanie: czy maszyny mogą myśleć? jest nazbyt nieokreślone, aby dać jednoznaczną odpowiedź. Zaproponował jednak słynny test na inteligencję, określany czasami jako „gra w udawanie". Test ten wywołał ożywioną i wielowątkową dyskusję nad użytecznością TM jako modelu ludzkiego umysłu i, ogólniej, nad możliwością stosowania maszynowych modeli w teoriach umysłu. Praca Piesko z pewnością zyskałaby na wartości, gdyby włączyła się w ową dyskusję, chociażby poprzez zasygnalizowanie tez najbardziej interesujących i(lub) problematycznych, dotyczących właśnie testu Turinga.

Dokładną analizę jednego z kluczowych zagadnień w problematyce AI – zagadnienia obliczalności – Autorka traktuje jako przyczynek do dyskusji nad AI, i dlatego nie tyle udziela odpowiedzi, ile raczej stawia pytania, nie tyle rozwiązuje problemy, ile raczej je uświadamia. Wydaje się, że książka ma za zadanie zorientować czytelnika odnośnie do teorio-obliczeniowego aspektu badań nad AI oraz spowodować, by próbował on szukać rozwiązań na wielu płaszczyznach (m.in. neurofizjologicznej, syntaktycznej, semantycznej). Na uwagę zasługuje ogromna staranność, z jaką Autorka kreśli genezę teorii obliczalności. Pod tym kątem dokonała ona, na kilkudziesięciu stronicach, syntezy najważniejszych osiągnięć (meta)matematyki od końca XIX w. Podkreślić należy także wartość przeprowadzonej przez Piesko krytycznej analizy pracy Turinga *On Computable Numbers...* i sformułowanie na jej podstawie trafnych wniosków.

Pobieżny rzut oka na spis literatury wykorzystanej przy pisaniu *Nieobliczalnej obliczalności* budzi pewne pytania. Chodzi przede wszystkim o nieobecność w nim klasycznych już pozycji z zakresu podejmowanej problematyki, takich autorów, jak np. J. D. Bolter (*Człowiek Turinga. Kultura Zachodu w wieku komputera*, Warszawa 1990), a na rodzimym gruncie M. Hetmański (*Umysł a maszyny. Krytyka obliczeniowej teorii umysłu*, Lublin 2000), S. Judycki (*Świadomość i pamięć*,

Lublin 2004), U. Żegleń (*Filozofia umysłu*, Toruń 2003). Można ją po części usprawiedliwić tym, że Piesko podjęła inny aspekt badanej problematyki. Czasami, jak się wydaje, pominęła pewne pozycje dlatego, że nie zgadzała się z niektórymi poglądami w nich zawartymi. Szkoda, że w swoich rozważaniach Autorka nie zrobiła „kroku dalej" i nie podjęła – wielce interesującej, jak sądzę – dyskusji np. z Hetmańskiego lub Judyckiego krytyką komputacyjnej koncepcji umysłu. Przeciwnicy komputacjonizmu uważają mianowicie, że tylko nieliczne działania poznawcze i praktyczne człowieka można opisać (modelować i symulować maszynowo) w kategoriach funkcji rekurencyjnych, obliczalnych. Natomiast swoistą „regułą" działania człowieka jest raczej to, że nie podlega ono ani wyłącznie, ani też najczęściej regułom dającym się ściśle opisać i obliczyć. W obrębie poznania i działania człowieka podkreślają takie cechy, jak niemechaniczność, niealgorytmiczność, nieobliczalność i niezdeterminowanie.

Lektura książki nasuwa jeszcze jeden wniosek: wszelkie dyskusje nad możliwością maszynowego modelu umysłu prowadzone w oderwaniu od koncepcji człowieka, czyli nie wchodzące na teren antropologii, wydają się jałowe. Nie jest to bynajmniej zarzut wobec Autorki, że tego aspektu nie ujęła. Książkę Piesko można (i należy) traktować jako przyczynek do dalszych interdyscyplinarnych badań. Całości problematyki AI nie da się zamknąć w jednej pracy, co więcej, nie da się chyba ogarnąć przez jednego człowieka.

Pomimo wyżej zasygnalizowanych braków, *Nieobliczalna obliczalność* wypełnia niewątpliwie pewną lukę w literaturze z zakresu AI czy szeroko rozumianej kognitywistyki. Odsłania w sposób dobitny, że w badaniach nad AI jesteśmy nadal na początku drogi. Celem publikacji nie jest rozwiązanie konkretnych zadań informatyczno-inżynierskich, prowadzących do konstrukcji „sztucznego mózgu", ani też dociekanie tego, czym jest inteligencja i jaka relacja miałaby zachodzić między inteligencją naturalną a jej sztucznym odpowiednikiem. Pole dociekań recenzowanej książki jest mocno zawężone. Sprowadza się ono do filozoficznej analizy teorii obliczalności, wywodzącej się od Turinga, w kontekście jej aplikacji do opisu działania programów komputero-

wych – potencjalnych „nośników" AI. Treść książki można by umiejscowić „pomiędzy obliczalnością a sztuczną inteligencją".

Nieobliczalną obliczalność warto polecić szczególnie filozofom, zaznajomionym (choćby pobieżnie) ze współczesną logiką. Zainteresować się nią powinni także przedstawiciele nauk formalnych – matematycy, informatycy, a także wszyscy ci, którym „smakuje" wysoki poziom abstrakcji prowadzonych analiz.

Bożena Czernecka-Rej
Wydział Filozofii KUL, Lublin